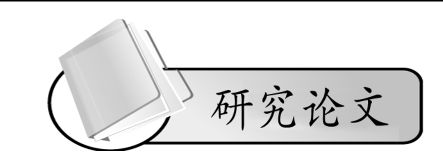


评价知识本体研究与规则实现*

周红照　侯　敏　滕永林
(中国传媒大学国家语言资源监测与研究有声媒体中心　北京 100024)

摘要:【目的】建立一套集评价句识别、倾向性判定、评价对象抽取于一体的评价分析智能系统。【方法】对中文评价知识本体进行研究，基于本体研究成果建立评价分析规则库，实现基于词典规则的评价分析智能系统 CUCsas。【结果】以第 7 届中文倾向性分析评测(COAE2015)发布的 50 000 篇微博(共计 133 201 个句子)为实验数据，系统评价句识别及倾向性判定的正确率、召回率和 F 值分别为 0.83、0.70、0.76，而评价对象抽取的结果较差。【局限】系统尚缺少评价新词发现和领域词典自动构建模块。【结论】初步建立起一套实用化的评价分析智能系统。

关键词: 评价知识本体　评价句识别　倾向性判定　规则方法　评价对象抽取
分类号: G350　H146

1　现有研究述评

评价指人们基于一定的标准对事物做出的褒义、贬义或褒贬混合的倾向性判断。评价分析(也称倾向性分析)近几年持续成为自然语言处理领域研究的热点问题之一，它在竞选预测、用户推荐、舆情监测、文献声誉追踪等诸多领域具有应用价值。

评价句(Evaluation Sentences, ES)包含 4 个基本要素：评价主体(Evaluation Subjects, Sub)、评价对象(Evaluation Objects, Obj)、评价因子(Evaluation Factors, Fac)、评价情景(Evaluation Context, Con)，即：ES = {Sub, Obj, Fac, Con}。评价分析智能系统(Evaluation Analysis Intelligent System, EAIS)指能够从混合文本中自动识别出带有评价意义的句子、判断句子的倾向性以及抽取句子中的评价对象的智能系统，通常包括评价知识本体(Evaluation Ontology, EOntology)与评价求解算法(Evaluation Solving Methods, ESM)两个基本模块，即: EAIS = EOntology + ESM。其中, EOntology 属于语言知识本体，指系统完成评价分析任务需要的语言知识，包括词汇知识(评价词典、否定词典、程度词典、转折词典等)、句法知识(词性、NP 和 VP 短语、依存关系等)以及语义知识(评价词语的语义指向特征、施事和受事语义角色等); ESM 有基于统计和基于规则两种基本方法。

(1) 基于统计的方法：张清亮等基于大规模语料，计算候选情感词与正、负情感种子词集之间的互信息(PMI-IR)，完成情感新词的识别和倾向性判定[1]；丁晟春等选取情感词、程度副词、否定词等特征训练 SVM 分类器，调整参数获取效果最优的模型完成句子倾向性分类[2]；童毅轩等采用集成学习的方法，将基于规则、基于 CRF 两种分类器的分类结果作为特征训练 SVM，进行微博观点句倾向性判定[3]；姜姗姗等选取词、位置、情感词、词性、父节点位置、依存关系、最近名词、基本短语类型、与情感词依赖关系、候选评价对象、特征词、关键词、句子极性、观点句共计 14 类特征训练 CRF 模型，抽取微博观点句中的评价对象[4]。

(2) 基于规则的方法：王昌厚等采用基于模式的自扩展方法(Bootstrapping)获取微博语料中的情感新

通讯作者：周红照，ORCID：0000-0001-6846-9006，E-mail：zhzwin2011@163.com。
*本文系国家语言资源监测与研究中心项目"中文情感倾向本体研究与规则实现"(项目编号: YZYS15-05)的研究成果之一。

词，例如“杭州/很/美丽/吧”，“很<instance>吧”就是一个模式，选择可信度高的模式匹配语料获取新的instance，如此反复迭代[5]；侯敏等通过制定情感短语计算规则，对“程度副词|否定词+情感词”的情况进行情感强度的增强、减弱和倾向性的翻转处理，提升句子倾向性判定的准确性[6]；周红照等基于语义分类构建包含“规则群–规则簇–规则”三个语义层级的比较句识别与比较要素抽取规则库[7]，采用句法分类和语义特征相融合的方法构建评价对象抽取规则库，实验效果良好[8]。

统计的方法基于经验主义哲学，认为语言知识蕴藏在语料中，机器通过对大规模语料的学习，可以自动挖掘语言知识；规则的方法基于理性主义哲学，认为语言知识(语法和语义规则等)蕴藏在人的大脑中，人们正是依靠存储在大脑中的有限规则生成和理解现实语言生活中的无限具体句子。这两种研究理念其实并不矛盾，只不过是分别从实例(Token)、类型(Type)两个不同的视角来看待“语言知识”问题。更进一步说，两者实质上是相互贯通的，“经验中蕴含理性，理性是对经验的抽象概括”。统计方法的概率计算、特征选取、模型参数训练等的目的就是使系统最大限度地贴近人们使用语言的规律，而且越是贴近语言规律的系统，性能表现通常也会越好。规则方法也从来不是“拍脑袋”，它从来都不拒绝语料和从语料中学习[9]，规则库中的每一条规则都不是凭空产生的，而是源自规则制定者自己头脑中储存的或他人表现出来的语言经验(真实语料)，规则制定者所做的工作是对具体语言经验的抽象概括，即将符合相同语言模式的大量实例抽象概括为一条规则，然后将规则组织成一个有机体系(规则库)处理类型尽可能广泛的语言实例。因此，统计和规则的不同只是所走的技术路线不同，两者并不存在本质差异。

无论是基于统计还是基于规则的评价分析方法，实质都是在探寻完成某项评价分析任务需要用到哪些基本的语言知识，以及如何将不同类型的知识有机组织起来。本文采用的是规则方法，规则的基本形式是“条件→动作”偶对，因此，面向自然语言处理的语言研究要着重阐明规则的条件[10]。本文首先从计算语言学的视角对中文评价知识本体进行研究，然后基于本体研究成果构建评价分析规则库，最后程序实现为基于词典规则的评价分析智能系统CUCsas，在第7届中文倾向性分析评测(COAE2015)中，该系统评价句识别及倾向性判定的准确率达到0.83，评价对象抽取的结果较差。

2 评价知识本体研究

评价分析是伴随互联网技术的发展(特别是进入Web2.0以后，博客、贴吧、论坛、微博、购物网站用户评论板块等社交媒体的兴起)，才开始成为自然语言处理领域关注的对象。而在此之前，Lyons[11]、Martin等[12]、Halliday[13]这些语言学者已做过“评价”的相关研究，其研究成果概述如下:

(1) 评价是语言主观性的类型之一，所表达的是一种主观性意义;

(2) 评价可以由韵律、形态、词汇、句法或肢体语言等手段体现;

(3) 评价在不同的说话语境中有不同的表达方式;

(4) 评价表达方式的选择主要由情景语境中的语旨要素决定;

(5) 评价实现的是语言的人际元功能;

(6) 评价是一种心理过程;

(7) 评价不同于表达对命题或建议的不确定性的情态;

(8) 评价可以分为判断与鉴赏;

(9) 评价有正、负倾向性之分，强、弱情感程度之别;

(10) 评价主体可以是说话人，也可以是由说话人转引的他人。

这些结论主要是从理论语言学的研究视角得出的，而计算语言学在研究对象、研究内容和研究方法方面具有不同于理论语言学的特点和要求。

2.1 研究对象：微观、具体

理论语言学围绕构建一套严密的理论体系展开，需要通盘考虑系统各个组成部分，并将其有机整合起来，研究对象宏观、多样。计算语言学的语言研究则是围绕某一具体的语言工程展开，它只聚焦于完成该项语言工程所需的那一部分语言知识，对其做深入细致的分析，研究对象微观、具体。

当人们从事评价分析这一语言工程时，首先要明确的是何谓“评价”？侯敏等把语句中表达评价意义的元素(包括评价词、评价短语、评价表达式、评价性句式4种颗粒度大小不同的语言单位)统称为评价因子[6]。确切地说，评价因子是指评价主体(说话人或说话人援

引的他人)在某一话语情景下(时间、空间、交际场景等)基于一定的评价标准(道德、审美、价值标准等)对某一评价对象(可以是任何事物)做出的带有主观倾向性(褒义、贬义、褒贬混合)的言语判断(词、短语、表达式、句式、语篇等)。褒贬倾向性是评价因子区别于其他概念因子的本质属性。“高兴、快乐、难过、悲伤”等情绪因子,“地震、战争、疾病、失恋、坠毁、遇难、改革、中奖、升官、订婚”等本身带有正面或负面内涵意义的事物、现象或事件因子,“韧性、气质、修养、才能、毅力、信用、创造力、流畅度、抗风险能力”等表示人和事物的品质属性的因子,均不具有独立表达褒贬倾向性的能力,因此不能算作评价因子。评价因子概念特征如表 1 所示:

表 1　评价因子概念特征

概念名称	概念特征
评价因子	[+主观情感, +外界刺激造成, +褒贬倾向]
情绪因子	[+主观情感, +外界刺激造成, –褒贬倾向]
评价因子	[+主观范畴, +内心世界, +概念意义是褒贬倾向, +没有内涵意义]
本身带有正面或负面内涵意义的因子	[+客观范畴, +外部世界, +概念意义是客观指称, +内涵意义是褒贬倾向]
评价因子	[+源自说话人, +主观依赖性, +形容词、动词为主, +褒贬倾向, –主观期望性]
品质属性因子	[+源自事物本身, +客观独立性, +抽象名词, –褒贬倾向, +主观期望性]

2.2　研究内容: 语言特征及其组合关系

规则的基本形式是“条件→动作”偶对,“条件”指的是语言特征及其组合关系。与理论语言学相比,计算语言学的语言研究主要侧重对语言现象的描写而非解释,它需要准确刻画出所研究的语言问题的客观状态,无需加以主观解释为什么会呈现这样的状态。因此,计算语言学很少像理论语言学那样考虑语言之外的社会、文化、心理等外部因素。不过,这并不意味着计算语言学的语言研究比理论语言学简单,理论研究重在思辨,对同一个问题可以有不同的解释,答案具有开放性、不唯一性;计算语言学重在实践,只是自圆其说不行,只有经得住大规模真实语料的检验才算有效,最佳答案通常只有一个,即性能表现最好的系统采取的问题解决方案。有些语言工程(例如机器翻译、人机对话)比较复杂,要对所研究的语言问题做出如实准确、面面俱到的描写并非易事,所以其研究周期会非常漫长,常常要耗费几代人的努力。

(1) 区别评价句、非评价句依据的特征

评价句包含评价主体、评价对象、评价因子、评价情景 4 个基本要素。其中,评价因子是评价句的必备要素,另外三个要素不强制出现。评价因子又分为上下文无关型和上下文相关型两种,前者不依赖语境,本身就表达评价意义,后者只有在特定上下文语境中才表达评价意义[8]。评价因子的语义类型如表 2 所示:

表 2　评价因子的语义类型

语义类型		示　例
上下文无关型		美丽、优秀、敬佩
上下文相关型	褒义性名词	客观、阳光、青春
	贬义性名词	主观、机械、封建
	度量衡形容词	高–低、大–小、长–短
	语义偏移型名词	素质、价值、水平
	语义构式	起到……作用、造成……问题、无法和……相比

但是,含有评价因子的句子未必就是评价句,例如:

①不管做什么事,只要勇敢地迈出第一步,成功的机会就有百分之五十。

②为了未来的幸福,好好奋斗吧!

③这是我送你的生日礼物,喜欢吗?

“勇敢、幸福、成功、奋斗、喜欢”都是评价因子,但在上述例句中并不表达评价意义。这是因为它们分别受到表示假设的连词“只要”、表示目的的介词“为了”、表示询问的句式“……吗”的管辖。笔者将这种具有取消其管辖范围内的评价因子的评价意义功能的词语和句式称为评价消解因子(Xjc),并将其归纳概括为以下类型,如表 3 所示:

表 3　评价消解因子的语义类型

语义类型	示　例
意愿	但愿、希望、祝、祝愿、祝福、祈祷
假设	如果、假如、若、倘若、假若、假设、要是、要想、即使、即便、就算、只要、当……时、……的话
猜测	猜、估计、无法判断、很难判断、很难说、能否、是否、会不会、有没有、是…还是…
询问	哪、哪儿、哪些、哪里、哪个、如何、为何、谁、为啥、为什么、什么、吗、么、呢、?
建议	应该、必须、应当、需、需要、请
商讨	讨论、探讨、商议、商量、商讨
目的	为、为了、以期、以便、目的是

因此,句子包含评价因子,且评价因子没有受到评价消解因子的管辖,才构成判定句子是评价句的充分条件,即: Fac + !Xjc => ED。

(2) 抽取评价对象依据的特征

语义指向是指句子中的词语在语义平面上支配或说明的方向[14]，评价因子的语义指向对象就是所要抽取的评价对象。抽取评价对象需要综合考虑评价因子的语义指向特征以及整个句子的句法结构特征来确定评价对象相对于评价因子的句法位置，继而抽取该位置上的成分(用词性或词性组合来表示)。语言中的具体句子是无限的，但作为构成句子基本单位的词语数量是相对有限的，词语的语法类型(词性)和语义类型也是相对有限的，词性、词义的组合方式(句法结构、句义结构)也并不是任意的，而是遵循一定的组合规律。因此，可以利用词形、词性、词义三个最基本的语言特征对评价各要素(Sub，Obj，Fac，Con)及其组合关系(包括 Fac-Obj 的语义指向关系)进行描写。与确定评价对象位置相关的语义特征主要有以下 4 类[8]如表 4 所示：

表 4　确定评价对象位置相关的语义特征

语义特征	内　涵	特征词示例
后指动词	评价对象通常是其后面的宾语。	佩服、鄙视、力挺、得益于
前指动词	评价对象通常是其前面的主语。	欺骗、污蔑、亵渎、有利于
心理动词	评价对象通常是其前面介词的宾语。	不满、愤慨、反感、爱不释手
前定名词	评价对象通常是其前面的定语。	功劳、奴性、毛病、凝聚力

这 4 类语义特征是基于语义指向方向对评价因子进行的细分，其功能是辨别同一句法结构具有两种不同的潜在语义结构的情况(形同义异)，例如：

④ a.孩子他爸佩服孩子他妈。b.孩子他爸欺骗孩子他妈。

⑤ a.孩子他爸对孩子他妈不满。b.孩子他爸对孩子他妈不好。

⑥ a.孩子他爸是孩子他妈的偶像。b.孩子他爸是孩子他妈的功劳。

上述例句中对举的两个句子语法结构完全相同，但评价对象(双下划线内容)却分布在不同的句法位置，要想对其加以区分，单纯依靠词形与词性特征显然无法完成。评价对象的位置分布之所以出现差异，是由评价因子(单下划线内容)语义指向特征的不同导致的，只有引入表 4 中的几类语义特征，才能准确判定这种“形同义异”句型评价对象的位置分布。

对于非“形同义异”句型，确定评价对象的位置主要是根据特征词(词形特征)与句法结构(词性组合特征)。笔者将与确定评价对象位置相关的词形和词性组合特征总结概括如表 5 所示：

表 5　确定评价对象位置的词形和词性组合特征

<table>
<tr><th>一级范畴</th><th>二级范畴</th><th>三级范畴</th></tr>
<tr><td rowspan="17">比较句</td><td rowspan="11">差比句</td><td>X＋没 | 没有 | 木有＋Y＋Fac</td></tr>
<tr><td>X＋不能 | 无法 | 没法 | 很难 | 难以 | 不够 | 不可能＋和 | 跟 | 与 | 同＋Y＋比 | 相比 | 相比较 | 抗衡 | 竞争 | 对抗 | 匹敌|同日而语 | 相提并论</td></tr>
<tr><td>X＋和 | 跟 | 与 | 同＋Y＋比|相比|比较|对比|相比较|比起来＋Fac</td></tr>
<tr><td>X＋和 | 跟 | 与 | 同＋Y＋没法比①</td></tr>
<tr><td>X＋连＋Y＋不如②; X＋不如＋Y</td></tr>
<tr><td>X＋好于③＋Y</td></tr>
<tr><td>X+是+Y+没法比④</td></tr>
<tr><td>X+比|较|相对|相比|相较|对比|比较|比起|相比较|相对于+Y+Fac</td></tr>
<tr><td>相比|相对|对比|比较|相比较|相较|相对于|比起+X, Y+Fac</td></tr>
<tr><td>X+在+Y+之上; X+不是+Y+对手|敌手</td></tr>
<tr><td style="display:none"></td></tr>
<tr><td rowspan="3">平比句</td><td>X+能|能够|足够|可以|足以+和|跟|与|同+Y+比|相比|比较|相比较|相提并论|抗衡|竞争|对抗|比肩|匹敌</td></tr>
<tr><td>X+、|和|跟|与|及|以及|或|还有|或者+Y+类似|一样|都+Fac</td></tr>
<tr><td>X+不输|不次于|媲美|匹敌|不逊色|不逊于⑤+Y</td></tr>
<tr><td rowspan="3">极比句</td><td>X(+在|是|算|有|拥有|具有|成为+中|里|里面)+最+Fac||首屈一指|无人能敌⑥</td></tr>
<tr><td>(在+中|里|里面)+最+Fac||首屈一指|无人能敌(+的+是)+X</td></tr>
<tr><td style="display:none"></td></tr>
<tr><td rowspan="6">非比较句</td><td rowspan="3">单句(单一结构)</td><td>主谓：主语[名词语|谓词语]+Fac</td></tr>
<tr><td>定中：Fac(+的)+定语中心语[名词语|谓词语];定语[名词语|谓词语]+的+Fac</td></tr>
<tr><td>动宾：动词性 Fac+宾语[名词语|谓词语]</td></tr>
<tr><td rowspan="2">单句(复合结构)</td><td>主语[名词语|谓词语](+状语[名词语|谓词语])+谓语中心语[动词]+Fac　定语[名词|形容词|动词]+宾语中心语[名词|谓词]</td></tr>
<tr><td>主语[名词语|谓词语]+Fac　状语[名词|形容词|动词]+谓语中心语[动词]+宾语[名词语|谓词语]</td></tr>
<tr><td>复句</td><td>分句 1+, |; |。|? |! (+主语承前或蒙后省略|照应语)+Fac(+宾语承前或蒙后省略|照应语)+, |; |。|? |! +分句 2</td></tr>
</table>

(说明：①②③④指 4 类差比词语簇，⑤指平比词语簇，⑥指极比词语簇。除了表中列举的词语，还包括与之同义的其他词语，例如差比词语簇①除了“没法比”，还包括“没得比、比不了、难以抗衡、有很大差距”等词语。)

之所以将评价句划分为比较句和非比较句两大基本范畴，是因为前者需要抽取的评价对象是复合型的，即：<比较主体 比较点>，<比较基准 比较点>，后者所抽取的评价对象是单纯型的，即：<评价对象属性>；在倾向性判定上两者也存在差异，例如，差比句的<比较主体 比较点>倾向性 = 评价因子倾向性，<比较基准 比较点>倾向性 = 评价因子倾向性×(–1)，{评价对象 属性}的倾向性 = 评价因子倾向性[15]。另外，从表 5 中可以看出，比较句评价对象位置的确定主要是根据特征词(词形特征)，非比较句评价对象位置的确定则主要是根据句法结构(词性组合特征)。

(3) 倾向性判定(褒、贬、褒贬混合)依据的特征

倾向性判定主要用到评价因子、否定词(否定副词、否定动词)、转折词三类特征，如表 6 所示：

表 6　倾向性判定依据的特征

特征类型	示例
评价因子	(见表 2)
否定副词	不、没有、毫无、尚未
否定动词	缺乏、欠缺、丧失、不足
转折词	但是、可是、然而、不过

若句中评价因子没有受到否定词管辖，Obj 倾向性 = Fac 倾向性。Fac 为褒义，Obj 为褒义；Fac 为贬义，Obj 为贬义。例如：

⑦这辆车性能很好。

⑧这辆车性能很差。

若句中评价因子受到否定词管辖，Obj 倾向性 = Fac 倾向性×(–1)。Fac 为褒义，Obj 为贬义；Fac 为贬义，Obj 为褒义。例如：

⑨这辆车性能不好。

⑩这辆车性能不差。

当倾向性相反的两个评价因子语义指向同一个评价对象(常用转折词连接)，Obj 倾向性 = 褒贬混合。例如：

⑪这辆车性能很好，但外观很丑。

2.3　研究方法：形式化、实验

理论语言学的研究目标是自然语言撰写的逻辑严谨的文章，研究方法主要是内省法。计算语言学的研究目标是程序语言编写的实用化系统，这就要求计算语言学的语言研究要采用机器可读的形式语言，并且在研究过程中不断进行实验，基于实验反馈情况对研究方法和结论进行修正。表 7 是笔者人工定义的评价知识本体形式符号集：

表 7　评价知识本体形式符号集

特征类型		形式符号示例
评价句识别及极性判定相关特征	上下文无关型评价因子	po(褒义) ne(贬义) cb (差比) pb (平比) jb (极比)
	上下文相关型评价因子	pxn(褒义性名词) nxn(贬义性名词) ypn(语义偏移型名词)
	评价消解因子	xjc(评价消解因子)
	正面程度副词	mopo(正面程度副词)
	否定词	mone(否定词)
	转折词	zzc(转折词)
评价对象抽取相关特征	词性	n(名词) v(动词) a(形容词) d(副词) r(代词) p(介词) c(连词) e(叹词)
	评价因子语义指向	hzv(后指动词) qzv(前指动词) xlv(心理动词) qdn(前定名词)
	评价对象性质	nq(产品名) na(产品属性名) bs(评价对象) ba(对象属性)

上述形式符号及其对应的具体词语以静态形式存储在词典中，词条示例：

[佩服 v]，[佩服 po]，[佩服 hzv]，[桑塔纳 nq]，[发动机 na]。

3　规则实现

利用表 7 中的形式符号，结合逻辑符号(或“|”、且“&”、非“!”)与运算符号(规则项分隔符“+”、通配符“*”、项位符“N_m”、赋值符“：”等)对第 2 节本体研究的语言特征及其组合关系进行描写，就构成了评价分析系统的规则库模块(问题求解算法 ESM)。评价分析规则库包括情感赋值规则库、情感计算规则库、评价对象抽取及极性判定规则库三个子库如表 8 所示。

规则的基本形式是条件→动作偶对，规则左部表示条件，规则右部表示在条件满足的情况下执行的操作。表 8 中示例规则的具体含义如下：

⑫ */mopo + */pxn = #2:0.75

含义：当褒义性名词(“青春、阳光、理性”等)受正面程度副词(“很、非常、特别”等)修饰时，赋予其 0.75 的情感值。

⑬ */xjc + #[*/!w] + */po|ne = #3:0

含义：当评价因子受评价消解因子(“但愿、如果、假如”等)管辖时，取消其评价意义。

⑭ */mone + */po|ne = -N2

表 8　评价分析规则库

类型		功能	规则示例
情感赋值规则库		识别上下文相关型评价因子并赋予其相应的情感值(上下文无关型评价因子的赋值已在情感词典中完成)。借助情感词典与情感赋值规则库，识别混合文本中的评价句。	*/mopo + */pxn = #2:0.75
情感计算规则库	评价消解规则	识别包含评价因子但并不表示评价的句子，消解掉句中评价因子的倾向性，对先前识别的评价句进行噪声过滤。	*/xjc + #[*/!w] + */po|ne = #3:0
	极性反转规则	完成倾向性判定：褒→贬，贬→褒。	*/mone + */po|ne = -N2
	褒贬混合规则	完成倾向性判定：褒贬混合。	*/po + #[*/!nq] + */zzc + #[*/!(w|nq)] + */ne = [N1 + N5]:2
评价对象抽取及极性判定规则库		抽取评价对象并判定评价对象的倾向性。	*/nq + #[*/!vl] + */na + #[*/!w] + */qzv = N1[bs] N3[ba] bs:N5

含义：当评价因子受否定词(“不、没有、毫无”等)管辖时，将其倾向性反转。

⑮ */po + #[*/!nq] + */zzc + #[*/!(w|nq)] + */ne = [N1 + N5]:2

含义：当一褒、一贬的两个评价因子语义指向同一个评价对象时，将评价对象的倾向性判定为褒贬混合。

⑯ */nq + #[*/!vl] + */na + #[*/!w|nq|na] + */qzv = N1[bs] N3[ba] bs:N5

含义：当产品名后面出现产品属性名，产品属性名后面出现评价因子且评价因子为语义前指动词时，抽取产品名、产品属性名为评价对象、对象属性，并将评价因子的情感值赋予评价对象。

需要说明的是，评价对象抽取及极性判定规则库中的规则并不是任意排列的，而是遵循一定的逻辑顺序。规则排列的基本原则是：比较句规则优先于非比较句规则，特殊句式规则优先于一般句式规则，复杂句规则优先于简单句规则。这样安排是为了让规则库尽可能模拟人脑的信息处理方式，例如：

⑰ Jeep在其广告中宣称的自由舒适，在操控性、乘坐空间等方面毫无体现。(错误抽取结果[Jeep　1];正确抽取结果[Jeep 操控性、乘坐空间　–1])

当规则库扫描到句子中出现产品名“Jeep”和评价因子“自由舒适”时，不会立即抽取产品名“Jeep”为评价对象，把评价因子“自由舒适”的情感值“1”赋予评价对象“Jeep”，而是会继续向后扫描，寻找是否存在产品属性名，是否存在管辖评价因子的否定词语。若存在“毫无体现”，则将评价因子“自由舒适”的情感值×(–1)，并且在抽取产品名“Jeep”为评价对象的同时抽取产品属性名“操控性、乘坐空间”为对象属性；若不存在，才会确认先前的判断正确。也就是说，规则库是按照最大匹配和确定性分析的原则来匹配语料的，这在一定程度上可以提高评价要素抽取的准确率。最后，将词典与评价分析规则库用 C#语言程序实现为评价分析智能系统 CUCsas。

4　实验结果与分析

CUCsas 系统参加了中国中文信息学会和中国计算机学会主办的第 7 届中文倾向性分析评测(COAE2015)，以评测发布的 50 000 篇微博(共计 133 201 个句子)作为实验数据，基于词典与规则的评价分析系统 CUCsas 的实验结果如表 9 所示：

表 9　评价分析系统 CUSsas 实验(评测)结果

任务	系统	微平均		
		正确率	召回率	F 值
评价句识别及极性分类	Best	0.8395	0.7851	0.8113
	CUCsas	0.8306	0.6984	0.7588
	Medium	0.6146	0.5642	0.6057
评价对象抽取及极性判定	Best	0.1117	0.2316	0.1508
	CUCsas	0.0495	0.2151	0.0805
	Medium	0.0684	0.1655	0.0866

实验(评测)结果表明，CUCsas 系统在评价句识别及极性分类任务上表现较好，尤其是正确率，达到83%，接近最好成绩(文献[3])；召回率稍差，在 70%左右。这证明了所选取的与评价句识别及极性分类相关的 6 类语义特征——上下文相关型评价因子、上下文无关型评价因子、评价消解因子、程度副词、否定词、转折词是合适有效的，情感赋值规则库和情感计算规则库对特征之间组合关系(规则条件)—组合结果(规则动作)的描写是基本准确的。评价消解因子、程度副词、否定词和转折词是规模较小、相对封闭的集合，评价因子则是规模较大、相对开放的集合，特别是在微博这样的自媒体领域，语言使用非常活跃，新兴评价因子不断涌现，而且夹杂许多口语化的表达，例如“坑

爹、伤不起、真特么尬”，由于这些评价词典没有收录评价因子，直接影响了系统的召回率，今后将加强词典未登录评价因子的自动发现研究。

评价对象抽取是整个评价分析中难度最大的一项任务，此次评测最好成绩(文献[4])的 F 值也只有 0.15，CUCsas 系统的表现更差，正确率和召回率都非常低。这一方面与此次评测任务难度较大有关，评测要求产品名、产品属性名、倾向性同时抽取正确才算正确，这在缺乏领域词典支持的情况下完成起来非常困难；另一方面，与评价对象抽取规则库还不够完善有关。今后将加强领域词典的自动构建技术研究，同时对当前规则库处理出错的情况以及没有覆盖到的情况进行修正与完善。

5 结　语

本文提出了一种面向评价分析(倾向性分析)的知识本体与规则库构建方法，首先从计算语言学的视角对完成评价分析任务(评价句识别、评价对象抽取、倾向性判定)需要的语言知识本体进行研究，然后基于本体研究成果建立评价分析规则库，最后程序实现为基于词典规则的评价分析智能系统 CUCsas。实验结果表明本方法具有较好的准确性。

下一步研究包括：探索词典未登录评价因子的自动发现机制，进一步提升系统评价句识别的召回率；加强领域词典自动构建技术研究，修正完善现有规则库，进一步提升系统评价对象抽取的性能。

参考文献：

[1] 张清亮，徐健．网络情感词自动识别方法研究[J]．现代图书情报技术，2011(10): 24-28. (Zhang Qingliang, Xu Jian. Research on Automatic Extraction of Web Sentiment Words [J]. New Technology of Library and Information Service, 2011(10): 24-28.)

[2] 丁晟春，李红梅．基于 SVM 的中文微博观点倾向性识别[C]．见：第 7 届中文倾向性分析评测会议(COAE2015)论文集.2015: 200-207. (Ding Shengchun, Li Hongmei. Chinese Micro-blogging Opinion Recognition Based on SVM Model [C]. In: Proceedings of the 7th Chinese Opinion Analysis Evaluation Conference. 2015: 200-207.)

[3] 童毅轩，张光磊．基于集成学习的中文微博情感分类研究[C]．见：第 7 届中文倾向性分析评测会议(COAE2015)论文集.2015: 55-61. (Tong Yixuan, Zhang Guanglei. Study of Chinese Microblog Sentiment Classification Based on Ensemble Learning [C]. In: Proceedings of the 7th Chinese Opinion Analysis Evaluation Conference. 2015: 55-61.)

[4] 姜姗姗，张佳师．中文评论细粒度情感分析[C]．见：第 7 届中文倾向性分析评测会议(COAE2015)论文集．2015: 70-78. (Jiang Shanshan, Zhang Jiashi. Fine-Grained Sentiment Analysis on Chinese Reviews [C]. In: Proceedings of the 7th Chinese Opinion Analysis Evaluation Conference. 2015: 70-78.)

[5] 王昌厚，王菲．使用基于模式的 Bootstrapping 方法抽取情感词[J]．计算机工程与应用，2014，50(1): 127-129. (Wang Changhou, Wang Fei. Extracting Sentiment Words Using Pattern Based Bootstrapping Method [J]. Computer Engineering and Applications, 2014, 50(1): 127-129.)

[6] 侯敏，滕永林，陈毓麒．评价短语的倾向性分析研究[J]．中文信息学报，2013，27(6): 103-109. (Hou Min, Teng Yonglin, Chen Yuqi. Research on Orientation Analysis of Opinion Phrases [J]. Journal of Chinese Information Processing, 2013, 27(6): 103-109.)

[7] 周红照，侯明午，侯敏，等．基于语义分类的比较句识别与比较要素抽取研究[J]．中文信息学报，2014，28(3): 136-141. (Zhou Hongzhao, Hou Mingwu, Hou Min, et al. Chinese Comparative Sentences Identification and Comparative Elements Extraction Based on Semantic Classification [J]. Journal of Chinese Information Processing, 2014, 28(3): 136-141.)

[8] 周红照，侯明午，颜彭莉，等．语义特征在评价对象抽取与极性判定中的作用[J]．北京大学学报：自然科学版，2014，50(1): 93-99. (Zhou Hongzhao, Hou Mingwu, Yan Pengli, et al. Function of Semantic Features in Opinion Target Extraction and Its Polarity Identification [J]. Acta Scientiarum Naturalium Universitatis Pekinensis, 2014, 50(1): 93-99.)

[9] 白硕．关于基于规则方法的反思[J]．心智与计算，2012，6(2): 66-74. (Bai Shuo. Ruminations on Rule-based Methods [J]. Mind and Computation, 2012, 6(2): 66-74.)

[10] 冯志伟．从知识本体谈自然语言处理的人文性[J]．语言文字应用, 2005(4): 100-107. (Feng Zhiwei. On Humanity Spirit of Natural Language Processing from the Viewpoint of Ontology [J]. Applied Linguistics, 2005(4): 100-107.)

[11] Lyons J. Linguistic Semantics: An Introduction [M]. Cambridge: Cambridge University Press, 1995: 293-342.

[12] Martin J R, White P R R. The Language of Evaluation:

Appraisal in English [M]. New York: Palgrave Macmillan, 2005: 1-91.
[13] Halliday M A K. Halliday's Introduction to Functional Grammar [M]. Revised by Matthiessen C M I M. The 4th Edition. London and New York: Routledge, 2014: 82-258.
[14] 范晓, 胡裕树. 有关语法研究三个平面的几个问题[J]. 中国语文, 1992(4): 272-278. (Fan Xiao, Hu Yushu. Some Questions about the Three Planes of Grammar Study [J]. Studies of the Chinese Language, 1992(4): 272-278.)
[15] 周红照. 基于句法语义的评价对象抽取与极性判定研究[D]. 北京: 中国传媒大学, 2014. (Zhou Hongzhao. Opinion Target Extraction and Its Polarity Identification Based on Syntax and Semantics [D]. Beijing: Communication University of China, 2014.)

作者贡献声明:

周红照: 研究设计, 论文撰写;
侯敏: 研究设计, 论文修改及定稿;
滕永林: 实验开展, 数据分析。

利益冲突声明:

所有作者声明不存在利益冲突关系。

支撑数据:

支撑数据由作者自存储, E-mail: zhzwin2011@163.com。
[1] 周红照, 侯敏, 滕永林. Eontology.rar. 评价知识本体库(包括分词词典、评价词典、语义词典等资源).
[2] 周红照, 侯敏, 滕永林. ESM_Rulebase.rar. 评价分析规则库(包括情感赋值规则库、情感计算规则库、评价对象抽取及极性判定规则库).
[3] 周红照, 侯敏, 滕永林. CUCBst.exe. 分词软件.
[4] 周红照, 侯敏, 滕永林. CUCsas.exe. 用 C#语言编写的评价分析程序.
[5] 周红照, 侯敏, 滕永林. COAE2015_Data.txt. COAE2015 评测数据(50 000 篇微博, 133 201 个句子).
[6] 周红照, 侯敏, 滕永林. Match_Result.txt. 实验结果数据.

收稿日期: 2016-05-16
收修改稿日期: 2016-07-01

Evaluation Ontology Research and Rule Implementation

Zhou Hongzhao Hou Min Teng Yonglin
(Broadcast Media Language Branch, National Language Resources Monitoring and Research Center, Communication University of China, Beijing 100024, China)

Abstract: [**Objective**] This study aims to build an evaluation analysis intelligent system consisting of evaluation sentences recognition, polarity identification and evaluation objects extraction. [**Methods**] We first researched Chinese evaluation Ontology. Then, established an evaluation analysis rule base based on the results of Ontology research. Finally, programmed into evaluation analysis intelligent system CUCsas. [**Results**] Taking 50,000 weibo messages (a total of 133,201 sentences) released by the 7th Chinese Opinion Analysis Evaluation Conference (COAE2015) as the experimental data, the precision, recall and F rates of evaluation sentences recognition and polarity identification of CUCsas were 0.83, 0.70 and 0.76 respectively, but the experimental result of evaluation objects extraction was poor. [**Limitations**] The system was short of new evaluation factors discovery and domain lexicons automatic construction modules. [**Conclusions**] A practical evaluation analysis intelligent system was basically built.

Keywords: Evaluation ontology　Evaluation sentences recognition　Polarity identification　Rule method　Evaluation objects extraction